

AWARD NUMBER: W81XWH-11-2-0167

TITLE: Framework for Smart Electronic Health Record- Linked Predictive Models to Optimize Care for Complex Digestive Diseases

PRINCIPAL INVESTIGATOR: John D Betteridge, M.D

Organization: The Henry M. Jackson Foundation for The Advancement of Military Medicine Inc.  
Bethesda, MD 20817-1834

REPORT DATE: March 2015

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) March 2015		2. REPORT TYPE Final		3. DATES COVERED (From - To) 1Jul2011 - 30Jun2014	
4. TITLE AND SUBTITLE  Framework for Smart Electronic Health Record- Linked Predictive Models to Optimize Care for Complex Digestive Diseases				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-11-2-0167	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)  John D Betteridge, MD  email: john.d.betteridge.mil@mail.mil				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)  The Henry M. Jackson Foundation for The Advancement of Military Medicine Inc.  6720 A Rockledge Drive Bethesda, MD 20817				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S Army Medical Research and Materiel Command Fort Detrick MD 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT  Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT We utilized de identified database constructed from the world-wide Military Electronic medical record for datasets (constructed for clinical purposes and not research) specific to patients with Crohn's disease (CD) and acute pancreatitis (AP). Using the collected data we developed Bayesian Network (BN) models which successfully predicted outcomes in CD and AP. We shared our model with the University of Pittsburgh Medical Center (UPMC) to confirm our modeling utilizing a different data set. UPMC also successfully predicted similar but not completely matching outcomes with their data set using the Bayesian Network Modeling.					
15. SUBJECT TERMS  Nothing Listed					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT  UU	18. NUMBER OF PAGES  32	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (include area code)

## Table of Contents

	<u>Page</u>
1. Introduction	4
2. Project Summary	4
3. Key Research Accomplishments	14
4. Conclusion	22
5. Publications, Abstracts, and Presentations	22
6. Reportable Outcomes	22
7. Appendices	24

## 1. INTRODUCTION:

Complex disorders result from the interaction of genetic, metabolic and environmental factors that may not produce disease themselves but combine to alter disease severity and its progression. These factors, which may be contained as relevant data in an EMR can be used to build predictive models with the hope of improving disease management.

It is difficult to find these factors in EMR systems as the information is stored in both structured and unstructured formats that have been collected over many years. Research studies in contrast only collect a limited snapshot of a patient's clinical history. This information is usually not rich enough to develop predictive models. To construct a useful patient profile requires collecting disease progression and treatment information from a wide variety of sources that may span a decade or more.

The University of Pittsburgh Medical Center (UPMC) has developed the Megascop application to provide a software platform for the integration of clinical, genomic and research data collected from multiple sources. The DoD, through its AHLTA, Essentris and CHCS electronic medical records and informatics systems has a worldwide data set of active duty, dependent and retired enrollees that is maintained in the Military Health System Data Repository (MDR).

Our study goal was to search our world-wide military medical record for data sets to develop Bayesian Network (BN) models to predict defined outcomes in Crohn's Disease (CD) and Acute Pancreatitis (AP) (death, surgeries, hospitalizations, etc.). We developed our data sets and cohorts then created predictive models that were validated using UPMC's Megascop model on our de-identified database.

The University of Pittsburgh's Department of Biomedical Informatics (DBMI), Division of Gastroenterology and Walter Reed National Military Medical Center (WRNMMC) Division of Gastroenterology, proved ideal collaborators given their expertise developing the informatics applications necessary and our clinical research in complex GI diseases and facility with the DoD electronic medical records systems.

## 2. OVERALL PROJECT SUMMARY:

As this study is a complex collaboration of partners inside and outside the DoD with unique expertise in their respective areas we will summarize the work by first identifying work done in direct collaboration with UPMC, WRNMMC, and the main DoD subcontractor Kennell and Associates. We will then specifically summarize the research findings which were significant enough for the WRNMMC group to present at national meetings over the past 3 years. Lastly we will identify key research methods that were essential in our project. As a collaborative effort, much of this work is overlapping and the summaries may be redundant in some respects where some or all parties are working on similar or identical problems.

### WRNMMC:

#### *JUL 2011- JUL 2012:*

This year, under direction of initial Primary Investigator (PI) Ganesh Veerappan was devoted to drafting, and gaining approval from Walter Reed, DoD Institutional Review Board for site specific protocol to perform the intended research as intended by the main study PI, Dr William A. Dunn. Also, CRADA agreements between all institutions and electronic data user agreements signed between Kennell Associates and Walter Reed. Initially, about 6 months was proposed timing necessary to gain appropriate IRB approval so that the award could be used for the planned research, but significant delays arose so that it took more about 15 months before local research could begin. This delay in anticipated

start date led directly to a need of no-cost extension of 1 year in 2013-2014. With future proposals for similar awards, especially those originating outside WRNMMC, I think that at least 1 year for administrative preparation prior to beginning research should be included in the timeline.

#### *JUL 2012- JUL 2013:*

Once IRB approval and agreed upon CRADA were in place in the summer of 2012, the planned research began. With our contracting agency, Kennel and Associates we defined AP case definitions, refined and organized the data set from the world-wide military database into a workable format for data software and statistical analysis. Beyond setting case definitions for AP we also set specific variables for extraction from the EMR (i.e. demographics, hospitalization, medication, ICU stay, and death) that might best allow for analysis of factors which can inform a clinician regarding outcomes with these chronic diseases. Further, we set these cases with de-identified data against a published hierarchical method of identifying the primary etiology for AP. In this way we were able to compare or extracted data against previous published outcomes in AP and CP. This served as useful validation of our data set before entering the Bayesian Network phase which began in the summer of 2013

Data analysis in AP began with a previously published method (Frey et al, Pancreas 2006) of hierarchical assignment of admitting etiology for each index case of AP. Thus using the index hospitalization, which presumably is an easily searchable event in any EMR one could better characterize the cohort at initial admission then create models based on re-admissions, other risk factors etc in predicting long term morbidity like recurrent AP, surgery, ERCP, chronic pancreatitis and mortality.

For CD, beginning in Spring of 2013 we followed a similar working plan and set case definitions and clinical variables for data extraction. Unlike the AP data set, hospitalizations in CD are uncommon so we could not use hospitalization for index case identification or primary outcome. Also, mortality in CD is very low so using death as a primary outcome measure to which we might map predictors of severity was impossible. Ultimately we settled on using Crohn's surgery as our primary outcome measure of severity and eventually built predictive models for ER visit and hospitalization as secondary outcome variables as our data set became more mature.

In order to validate the occurrence of Crohn's surgery we decided against simply using ICD-9 and CPT codes to identify those in the cohort with a surgery indicated by symptomatic Crohn's disease. We identified the cohort using coding but searched for a way to validate the coding to be more specific in our case and outcome definitions. Our colleagues at UPMC had proven a method within their own EMR of digitally mining their own EMR for surgical operative reports and were able to more reliably validate Crohn's specific surgery in their cohort. Our research was limited in two ways for applying a similar approach. First, digital text surgical operative reports do not exist in the MHS data repository. Second, our research subcontractor, Kennel and Associates did not have capabilities for mining NLM text concepts through a computer algorithm. To validate our desired outcome we tried a new source method, surgical pathology reports which indirectly could identify findings of Crohn's disease in nearly all Crohn's surgical specimens. After deciding to use surgical pathology reports as our source EMR document, we set up a subcontract to UPMC who validated a text concepts method of extracting data using pathology reports as a way to validate the occurrence of Crohn's surgery

#### *JUL 2013-JUL 2014:*

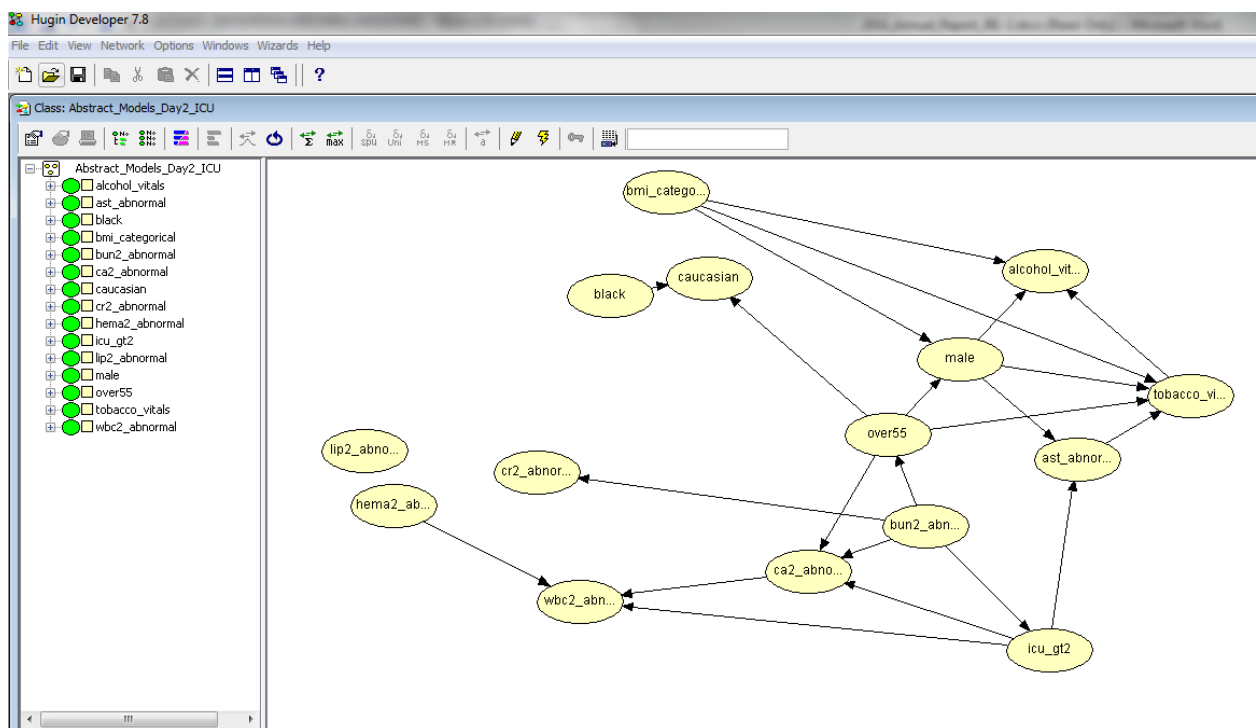
Our WR research team focused on working with our collaborators on the AP and CD cohorts to develop our BN predictive models. More specifically, creating and testing Bayesian Network models for AP and setting case definitions and extracting data for generating the CD data including testing a new text concepts method of digital text mining pathology reports to confirm our clinical outcome of Crohn's surgery. All this was in preparation of creating our predictive models for CD.

We met monthly and collaborated with our colleagues at Kennell and Associates to develop Bayesian Network predictive models for outcomes associated with admission episodes for AP. We successfully created a database generated solely from the MDR and used that wide array of data points to create Bayesian Network models of AP as a disease and identify significant variables which may predict morbidity as defined as organ failure or prolonged ICU or hospital stay. Building on our previous years of collaboration with our contracting agency, Kennell and Associates we had defined AP case definitions, refined and organized the data set from the world-wide military database into a workable format for data software and statistical analysis. We set these cases with de-identified data against a published hierarchical method of identifying the primary etiology for AP. In this way we were able to compare or extracted data against previous published outcomes in AP and Chronic Pancreatitis (CP). This served as useful validation of our data set before entering the Bayesian Network phase for this data set. Again, in close collaboration with our colleagues at Kennell and Associates we developed a dichotomized set of variables from our large AP data set with more complex, but still clinically relevant definitions which would allow for computer algorithmic assignment of contributing importance to outcomes as well as machine learning and validation of the predictive model.

## Bayesian Modeling for AP

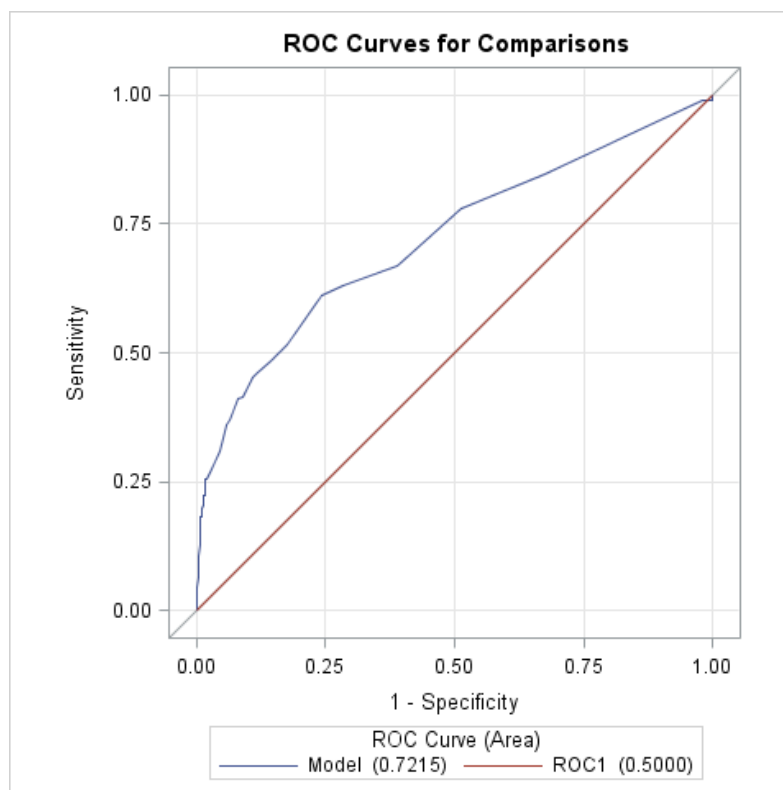
After obtaining the data for the AP cohort, binary variables were created and used for Bayesian modeling as well as Logistic Regression analysis. We created four models for each outcome variable during the AP admission (ICU stay greater than two days and any organ failure), which focused on demographics, abnormal laboratory results, and etiologies of the patient. For the Bayesian modeling, we used Hugin Developer, which used EM-Learning to create a Bayesian Network (Figure 1).

Figure 1: Bayesian Model (Day 2 labs) for ICU stay greater than two days (outcome)



For each model, the data was partitioned into five mutually exclusive random samples. Each sample was then paired with a 'match-sample' that contained the other 80% of the data. The 'match sample' was used for the EM-Learning in Hugin to build the Bayesian Network and the 20% was used to test it. The output for each of the five 20% samples provided us the predicted probability for every AP admission and was used to create a ROC Curve (Figure 2).

Figure 2: ROC Curve for Bayesian Model (Day 2 labs) for ICU stay greater than two days (outcome)



In summary, Hugin Bayesian Network Models were constructed to predict patients with severe acute pancreatitis (AP) defined as ICU hospital stay greater than 48hrs. Then Four (4) separate models were constructed using demographic (age>55yrs, Race, Gender, BMI, Tobacco use, Alcohol Use), lab values (BUN, Creatinine, Lipids, WBC, Hemaglobin, AST, Calcium) at initial admission and on Day of admission, and etiology of AP. Logistic Regression using the same set of variables was also performed to confirm the Bayesian model.

**MODEL #1** – using Demographic, and Lab values at initial admission:

The Hugin Model directly linked 4 variables to ICU >48hrs: abnormal AST, WBC, Calcium, and Chromium. This was confirmed by LR which identified the same variables as well as BUN.

**MODEL #2** – using Demographics, Lab values at initial admission, and etiology:

The Hugin Model linked the same four variables as in MODEL #1 to ICU > 48hrs. No additional improvement in the model was obtained by adding etiology.

**MODEL #3** - using Demographic, and Lab values on Day 2 of admission:

The Hugin Model directly linked 4 variables to ICU >48hrs: Ca, BUN, WBC and AST. This was confirmed by LR which identified the same variables as well as Cr2.

**MODEL #4** - using Demographics, Lab values on Day 2 of admission, and etiology:

The Hugin Model directly linked 3 variables to ICU >48hrs: Ca, BUN, and WBC. This was confirmed by LR which identified the same variables as well as Cr2.

***The most significant variables for predicting patients with severe AP defined as ICU stay greater than 48hrs, is an abnormal Ca, BUN, WBC and Ast at initial admission or on Day 2 of admission plus Cr on Day 2. Etiology of AP does not have a significant impact on severe AP.***

Lastly, we sent all eight models to the University of Pittsburgh for validation with their data. Each model was re-run in Hugin using all AP admissions (for the selected variables in each model) instead of just the 80%. We created a Bayesian Network for each model using the EM-Learning in Hugin and sent the model (not the data) to the University of Pittsburgh along with documentation.

For CD we have followed a similar working plan and set case definitions and clinical variables for data extraction. Kennell and Associates defined CD case definitions, refined and organized the data set from the world-wide military database into a workable format for data software and statistical analysis. Beyond setting case definitions for CD we also set specific variables for extraction from the EMR (i.e. demographics, hospitalization, medication (biologics, steroids), surgery, and death) that might best allow for analysis of factors which can inform a clinician regarding outcomes with these chronic diseases (see below). One delay in our work with the CD data set has been finding a reliable reproducible outcome. Given MHS available data in CD, we preferred to have a model of disease with the base unit being the patient rather than an index hospitalization as in AP.

CD for the most part is managed in the outpatient setting and even flares are managed as such, thus in collaboration with our colleagues at UPMC we settled on a primary outcome of Crohn's surgery. This is searchable in the MDR (CPT and ICD 9 codes) and also potentially verifiable in the EMR. The major obstacle here is that surgical operative notes are not widely searchable in the MHS EMR, thus we decided to use pathology specimen reports as a surrogate EMR source for the same outcome. In order to verify the outcome of Crohn's surgery, free text data mining of surgical (anatomic) pathology became necessary. We then set up a relevant Text concepts dictionary of terms for search in accordance with National Library of Medicine (NLM) published dictionaries. This work was sub-contracted to our colleagues at UPMC who had experience with such text mining in CD in their own EMR using surgical operative reports and is nearing conclusion (see below).

This is an exciting advance as it proves a new way of abstracting previously unusable data in the MDR, meaning dictated physician notes or encounters like surgical operative reports, radiology exams, pathology reports, and even clinical notes. Though currently unavailable within the MHS, the work done at UPMC was done utilizing mostly open source software could be invaluable for gathering and evaluating data available in text form in the DoD EMR. One drawback we encountered using pathology reports is that it was very specific but we think less sensitive than using surgical operative reports to validate a Crohn's specific surgery. Most, but not all Crohn's disease surgeries result in a pathologic specimen, but an incision and drainage of abdominal or pelvic abscess or fistula in some cases may not.

Due to the delays, specifically in starting this text mining research, a second no-cost extension for 31JUL 2014-31 DEC 2014 was sought and approved, see attached letter.

#### *1 AUG 2014- 31 DEC 2014:*

The last phase of research focused on our collaboration with Kennell and Associates and UPMC. Our Bayesian Network models were validated by the UPMC informatics group. The data set created by Kennell was finalized. Additional variables were added to the data set, specifically more specific medication use data including antidepressant use, biologic and steroid use and mood disorders within the cohort. We refined and created the outcome data set gleaned from the text mining work done by our UPMC colleagues. Finally the Kennell team with frequent interaction with the WR group loaded and tested our Bayesian Network modeling with our data set, creating three predictive models in CD, identifying variables that predicted Crohn's surgery, ED visits and hospitalizations. These models were referred to our colleagues at UPMC who validated the predictive models.



## **Collaboration with Kennell and Associates:**

The following is an overall summary of work done over the course of the study rather than broken into yearly periods as above. In general, Kennell worked one on one with WR research team to extract data from the MHS data repository and organize data sets and perform statistical experiments on data. Thus, effectively creating and validating the cohorts for AP and CD respectively. The Kennell team internally developed techniques of Bayesian Network modelling and successfully created our predictive models. The Kennell team also coordinated and oversaw secure data transfer of de-identified data between the DoD and our collaborators at UPMC. Details on individual tasks as per Kennell team follow below:

Task 1– Maintained data use agreement with the TMA Privacy and Civil Liberties Office (PCLO) to obtain access to the necessary data files

Task 2 – Developed functional specifications for analytic data files based on the case definitions and search parameters established by Dr. Betteridge and his team

Task 3– Write programming code

Task 4– Run programming code and conduct quality

Task 5– Where necessary, extract and evaluate text files and interpret text concepts to be included in the analytic data files

Task 6– Prepare data dictionary and any additional data documentation

Task 7– Assisted in the writing of any final reports or publications required for the study

Task 8– Attended annual project meetings with UPMC research team in Pittsburgh, PA

Task 9– Attended monthly project status meetings with Dr. Betteridge and his research team at WRNMMC, UPMC subcontracts and collaborates with WRNMMC Research Team and Kennell and Associates

For the AP and CD cohorts, the Kennell team developed functional specifications for analytic data files based on the case definitions and search parameters established by Dr. Betteridge and his team. The Kennell team wrote all programming code and conducted quality and validation checks. The Kennell team ran all programming code for data extraction and where necessary, extract and evaluate text files and interpret text concepts to be included in the analytic data files. Kennell and associates created the data dictionary for the data analytic reports and prepared the AP cohort for Bayesian Network Analysis and initiated and managed UPMC subcontracts and collaborate with WRNMMC Research Team and Kennell and Associates to interpret text concepts and apply Megascope to the MHS electronic health records of the study cohort. Specifically, the following tasks were performed by Kennell and Associates to meet the objectives.

### **For AP Cohort**

For Acute Pancreatitis, initial data extraction showed that Acute Pancreatitis accounts for around 7,000 hospitalizations to direct care DoD facilities of MHS beneficiaries annually. In identifying the AP cohort a unit of analysis was the AP admission to direct care Military Treatment Facilities (MTF) including Ft. Carson, Ft. Gordon, Tripler, Ft. Campbell, Walter Reed, Bethesda, Ft. Bragg, Ft. Bliss, Ft. Hood, Ft. Belvoir Community Hospital, Madigan, Landstuhl, Travis AFB, San Diego, NH Jacksonville, NH Pensacola, Camp Lejeune, Wright Patterson, Ft. Sam Houston, Lackland, Portsmouth. Utilizing ICD 9 codes (ICD9-9-CM-577.0) a data cohort was identified (see below).

<b>Number of Acute Pancreatitis Admissions</b>						
	FY 2009	FY 2010	FY 2011	FY 2012	FY 2013	<b>Total</b>
<b>ALL AP Admissions (Direct Care)</b>	835	799	776	857	153	<b>3428</b>
<b>AP Prime Admissions</b>	551	528	544	611	110	<b>2350</b>
<b>AP 65+ Admissions</b>	155	147	121	130	23	<b>578</b>

During analysis we found the military data set very similar to previously published regional cohorts for AP, except that our DoD data set at direct care MTF were significantly younger (see below). This likely reflects a limitation in searching the MDS for usable data which we have encountered with this project. Care received by Tricare beneficiaries “on the economy” is common. This is traceable by the Tricare claims data, however usable clinical data in generating predictive models requires access to laboratories, and radiographic studies, pharmacy records, even clinical notes. For this reason we limited our cohort to direct care MTF as above, which in turn likely gave us a younger cohort that is not reflective of the entire Tricare population.

Following creation of the research cohort we successfully applied a computer program written by Kennell team members to create a computer based algorithmic model for assigning each index case of AP an etiology for admission. This allowed for comparison not only based on demographic data and clinical parameters but a more detailed analysis by disease etiology over time when also applied to recurrent cases.

Data extraction for Crohns disease began in Summer 2013. Similar to AP, Kennell and Associates created a data set from patients receiving care at direct care MTFs using the following algorithm:

Unit of Analysis: Patient (identified at initial CD diagnosis)  
Qualifiers for Cohort:  
Two ICD-9-CM Diagnosis Codes (555.0-555.9) within 1 year  
One Diagnosis Code and prescription within 6 months  
Three Diagnosis Codes within 1.5 years  
Clean Period: 2 yrs. (prior to first diagnosis)  
Follow up Period: 3 yrs. First diagnosis in FY 2009 or 2010.s.

<b>MHS Wide</b>		
<b>Number of patients</b>	<b>FY 2009</b>	<b>FY 2010</b>
Crohn's diagnosis (all-listed)	17,822	18,826
Qualified for Cohort	1,190	1,125
Percentage	6.70%	6%

Direct Care Only		
Number of patients	FY 2009	FY 2010
Crohn's diagnosis (all-listed)	4,467	4,522
Qualified for Cohort	555	516
Percentage	12.40%	11.40%

With these defined cases we identified disease predictors which may predict things like intestinal surgery, requirement of ostomy, disability. Potential predictors included age, race, labs, Charlson Comorbidity Index, history of fistula, cancer, etc, disease location descriptors, Lupus, Thyroid disease, Rheumatoid arthritis, history of Appendectomy, psoriasis.

The Crohn's work is also unique in that we relied on our ability through our collaboration with UPMC to design and implement text reading software for validated NLM text concepts that were extracted from the EMR physician pathology reports. This work will validate our predictive models and hopefully open up a new pathway that other DoD researchers can use similar digital data extraction methods in a host of other diseases.

After Kennell provided UPMC with de-identified pathology reports from the MDR, UPMC was able to use the library of terms they developed based on their own surgical notes to extract those terms in the pathology reports that were positively associated with GI surgeries for Crohn's Disease. Kennell was able to link those results back to the original pathology reports, and further back to the surgical record to identify which procedure code (CPT or ICD-9) was listed for that surgery. By combining the two data sources, Kennell is able to find surgeries specifically related to Crohn's Disease, which was not possible without utilizing the text-based pathology reports. Furthermore, the text concepts pulled from the pathology reports can be further used to define the disease location, the anatomic location of the disease, and a classification of the biologic behavior.

Once the cohort of Crohn's surgeries was identified the Kennell and WR research teams concentrated the majority of efforts over the final 4 months of research work in from AUG to NOV 30 2014 on dichotomizing the data set and creating and testing a Bayesian Network predictive model for risk factors that would predict a Crohn's specific surgery. Multiple models were created with expansion and retraction of data criteria as follows:

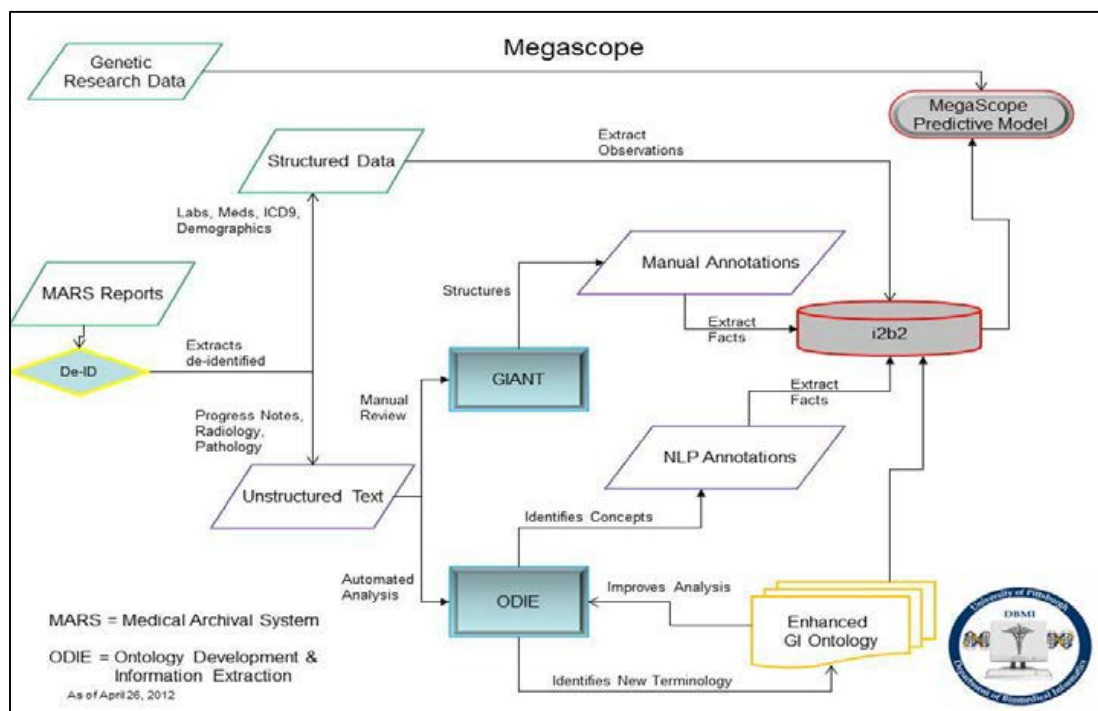
1. Data from cohort data set including validated Crohn's surgeries were entered into Hugin software
2. Multiple, (6) different models were entered and tested using multiple variables
3. Models for ED visits, Admissions, and Crohn's/GI surgery was completed and tested with resulting ROC curves and confidence intervals for each model were produced

For each model, the data was partitioned into five mutually exclusive random samples. Each sample was then paired with a 'match-sample' that contained the other 80% of the data. The 'match sample' was used for the EM-Learning in Hugin to build the Bayesian Network and the 20% was used to test it. The output for each of the five 20% samples provided us the predicted probability. Below is an example of completed model for GI surgery in CD and ROC curve with estimated predictive probability of the model.



## Work in Collaboration with UPMC:

Again, here is an overall summary of work rather than a yearly breakdown of our collaboration. Our work and meetings with University of Pittsburgh Medical Center focused around development of variables for algorithmic analysis for the application of Bayesian Network analysis. We have learned that the digital data sources we are using contain important clinical data in vastly different formats. UPMC has used their experience with Megascope (See figure below) to consult and guide the creation of workable data extraction programs applicable to MDR. In the second goal of creating the data cohorts themselves for AP and CD we are working to subcontract directly to UPMC.



The main thrust of our collaboration with UPMC in development of algorithmic variables for the IBD Cohort: UPMC identified the specific outcome of surgery for their Crohn's set by processing the operative reports through their "phenotyping pipeline". During previous meetings with UPMC colleagues, we alerted them that the operative notes are not searchable in the Military cohort and we needed to modify our approach to identifying surgical events. Our revised approach is to identify a surgical outcome via a surgical pathology report since pathology reports are available on the military cohort. Once contractual agreements were in place UPMC took the de-identified pathology reports provided by Kennell and Associates and applied their previously established NLM dictionary for CD they developed by processing operative reports from their own EMR in the UPMC system. In their experience with GI surgeries for CD they found that GI surgery domain is not well represented in standard ontologies. So, they are adding each of the procedure terms to our ontology and will contribute this ontology to the National Center for Biomedical Ontology ([www.bioontology.org](http://www.bioontology.org)) upon completion. This is where collaboration with UPMC was crucial. Since work like this has not been done with MHS data, their experience in using text mining and creating and validating text concept searchable terms which would correctly identify a pathology report from a CD specific surgery eliminated several developmental steps in applying this technology to the MHS data. Without this collaboration this work could not have been completed this year.

UPMC has developed a mature method for data extraction and disease characterization or "phenotyping" called GIANT, a web-based annotation tool enables researchers to annotate de-identified clinical reports. The application design focuses on providing users with an intelligent

workspace, by displaying annotation forms and de-identified reports with the same view, automatic report queuing and providing easy access to annotation guidelines and data definitions. The application produces user statistics to report agreement between multiple annotators who are reviewing the same report. This tool was built using the Django ([www.djangoproject.com](http://www.djangoproject.com)) web framework, which is an open-source project built on the Python ([www.python.org](http://www.python.org)) programming language. The annotation tool features include controlled user access, database support, progress reporting, task-specific error checking and a site administration interface. There are two output streams for GIANT. The first output is the report annotations completed by the clinical expert that will be imported into i2b2. The second output is the list of concepts identified in ODIE that appear most frequently in documents. This concept generator is used for feature selection to comprise the elements in the predictive model.

In order to search and derive variables from free text electronic sources like pathology reports, radiology or endoscopy reports a reliable system of ontology processing is required. The same condition exists with identifying acute pancreatitis (AP) in radiology reports. A possible future project with UPMC is development of specific ontology applicable to pathology and radiology reports within the DoD data sources. Since ODIE identifies both concepts (CUI) and semantic types (TUI) found in the narrative reports. These data will be used as the input for clinical variables in building the prediction models.

Walter Reed and UPMC have agreed to present and publish their individual research findings separately with regard to the clinical findings of the AP and CD research cohorts. It was also agreed, the Dr William Dunn's informatics team at UPMC take the lead on a combined manuscript publishing the data informatics findings and novel techniques utilized in this study. This work is ongoing at the time of this Final Report.

### **3. KEY RESEARCH ACCOMPLISHMENTS:**

Summary of Research Findings Presented at National Meetings by WRNMMC Research Team.

#### **Acute Pancreatitis:**

We presented our initial data collected on acute pancreatitis at Digestive Diseases Week in May 2013. This data was primarily an epidemiologic description of the burden of acute pancreatitis within a nationwide military population. We calculated an annual incidence (36/100,000) and characterized the demographics of the patients diagnosed with acute pancreatitis. Our calculated incidence was similar to those calculated from other published studies looking at nationwide populations. We identified rates of important outcomes associated with acute pancreatitis (death, ICU admission, organ failure, surgery). Using logistic regression, we identified risk factors that were associated with a severe course of acute pancreatitis (defined as AP resulting death, ICU stay >48hrs or organ failure). Our results indicated that age >55, AST>250, WBC >16,000, serum glucose >200, and Calcium <8 on admission were all associated with an increased likelihood of severe pancreatitis.



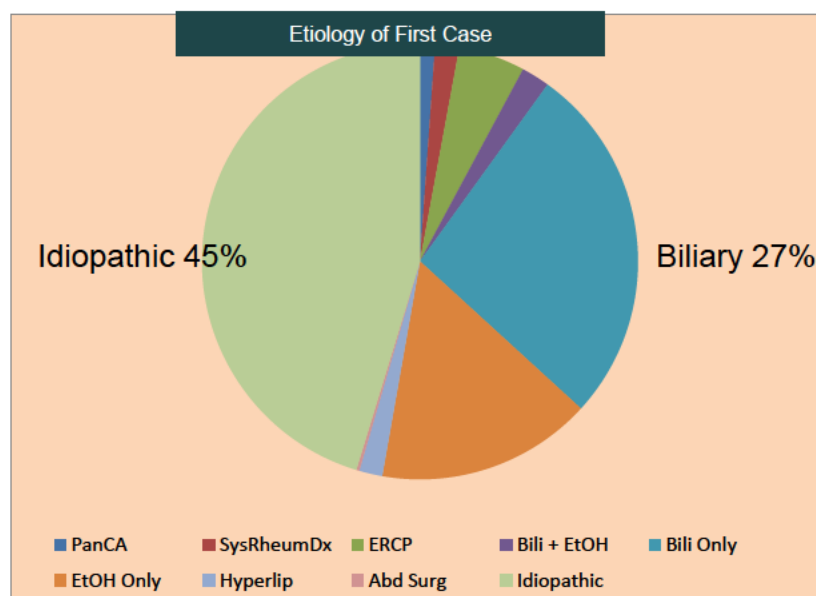
## Demographics

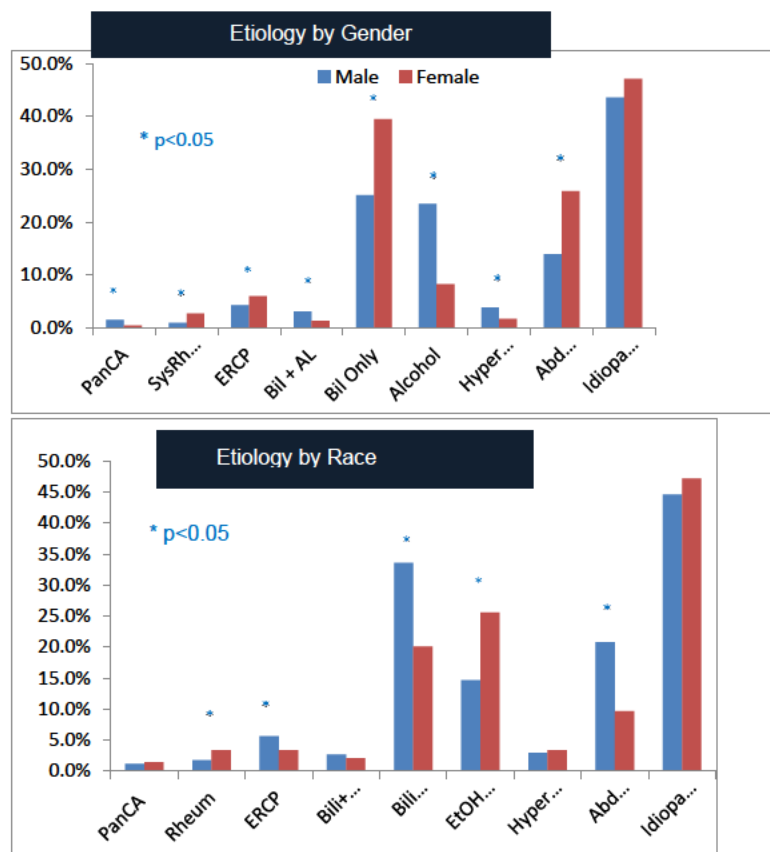
Pancreatitis cases (2008-2012)	1,767
Average number of cases/year	441
Male	48%
Race	
Caucasian	52%
African American	17%
Other	30%
Mean age (y $\pm$ SD)	41 $\pm$ 14
Total # hospital days (Mean $\pm$ SD)	4 $\pm$ 4.7
Total ICU days (Mean $\pm$ SD)	0.5 $\pm$ 0.27
Readmission rates	18.5%

## Predictors of Severe Pancreatitis on admission

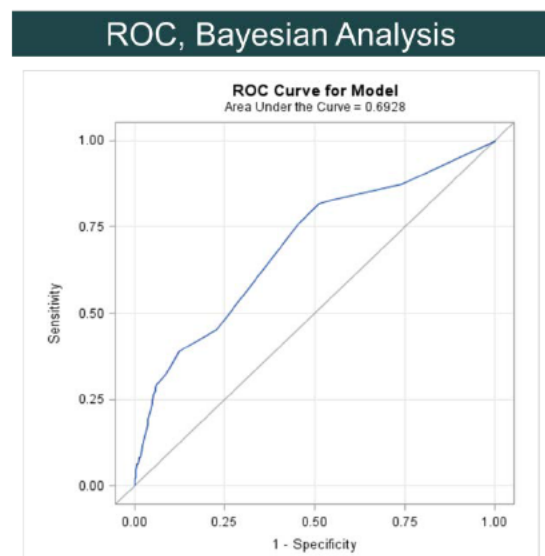
	Adjusted OR (95%CI)	P-value
Age > 55yrs	1.64 (1.01, 2.67)	0.044
AST>250	2.22(1.13,4.38)	0.02
WBC >16,000 cells/mm <sup>3</sup>	2.38(1.28, 4.41)	0.006
Serum Glucose >200mg/dL	2.25 (1.14, 4.46)	0.019
Serum Calcium <8 mg/dL	6.35 (3.15, 12.8)	0.000

In October 2013 we presented data which looked primarily at the differing etiologies for acute pancreatitis. This data set included an expanded patient population (patients >65 were included). We looked at how these etiologies varied with sex, age and race. We also examined those patients who had recurrent acute pancreatitis and identified which patients were more likely to have recurrent pancreatitis. We found that many cases of acute pancreatitis did not have a defined etiology (idiopathic). Forty-five of cases were idiopathic. The next most common cause of AP was biliary (27%), followed by alcohol (19%). Females were more likely to have biliary pancreatitis, while males were more likely to have alcoholic pancreatitis. Caucasians were more likely to have biliary pancreatitis, while Blacks were more likely to have alcoholic pancreatitis. Alcohol was found to be the most common etiology for recurrent AP. Idiopathic pancreatitis was the number one cause of AP among all different age cohorts.



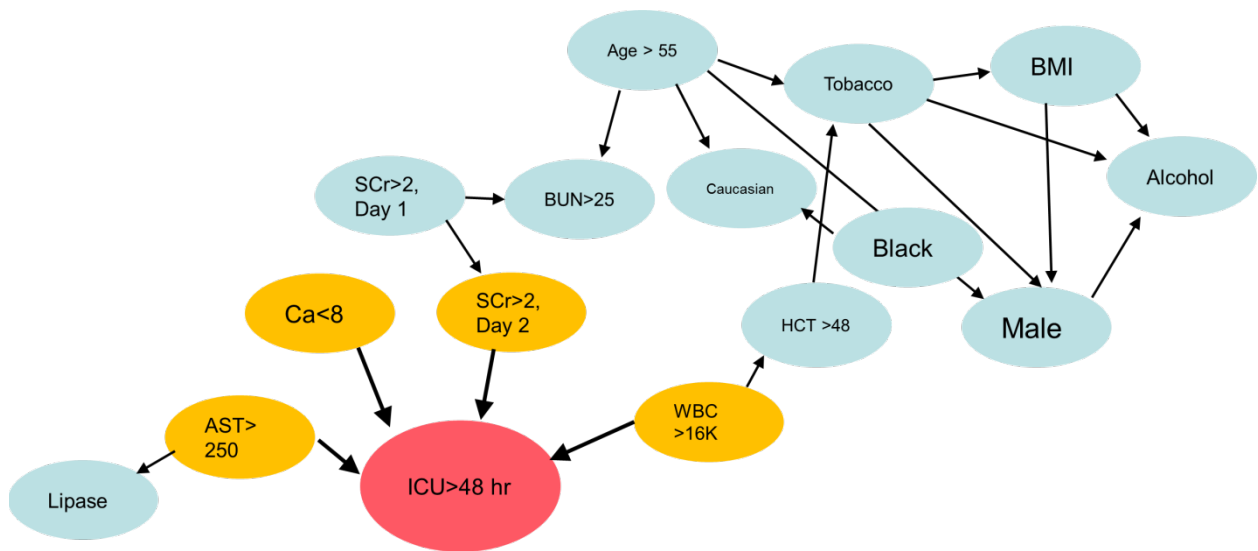


In October 2014, we presented data at the American College of Gastroenterology (ACG) annual meeting. This data included the results of the Bayesian network analysis for risk factors to predict severe acute pancreatitis. The Bayesian analysis identified the following as predictors of severe acute pancreatitis: AST>250, serum calcium < 8mg/dL, serum creatinine >2 on day 2, and white blood cell count of >16,000/mL. A separate logistic regression analysis was also performed which also identified the same variables as being the most predictive of severe acute pancreatitis. The ROC curve for the Bayesian analysis was modest (AUC 0.693), but the fact that there was agreement with the logistic regression analysis suggests that the method is a valid method of identifying variables in a broad data field that can be predictive of a defined outcome. We believe that results of this study lends credence to the notion of Bayesian analysis as a tool for predicting outcomes of various medical conditions and identifying risk factors in patients that may predict more severe outcomes.



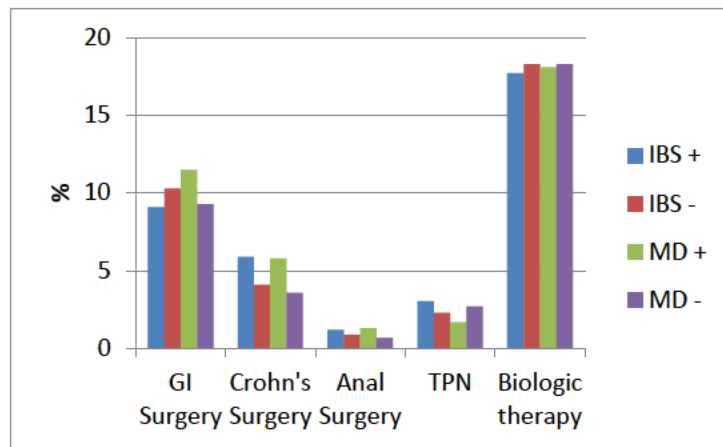
Logistic Regression	
Variable	Odds Ratio, 95% CI
Age > 55	1.226 (0.797-1.885), P 0.355
Race (White vs Af Am)	1.176 (0.702-1.971), P 0.537
Sex (Male)	1.368 (0.885-2.113), P 0.158
Hematocrit >48	1.779 (0.926-3.414), P 0.084
BUN > 25	1.962 (1.010-3.812), P 0.047
AST >250	2.480 (1.629-3.776), P <0.001
Serum creatinine > 2	8.057 (3.403 – 19.08), P<0.001
Serum calcium < 8	4.080 (2.040 – 8.16), P <0.001
WBC > 16, 000	2.886 (1.796-4.639), P <0.001



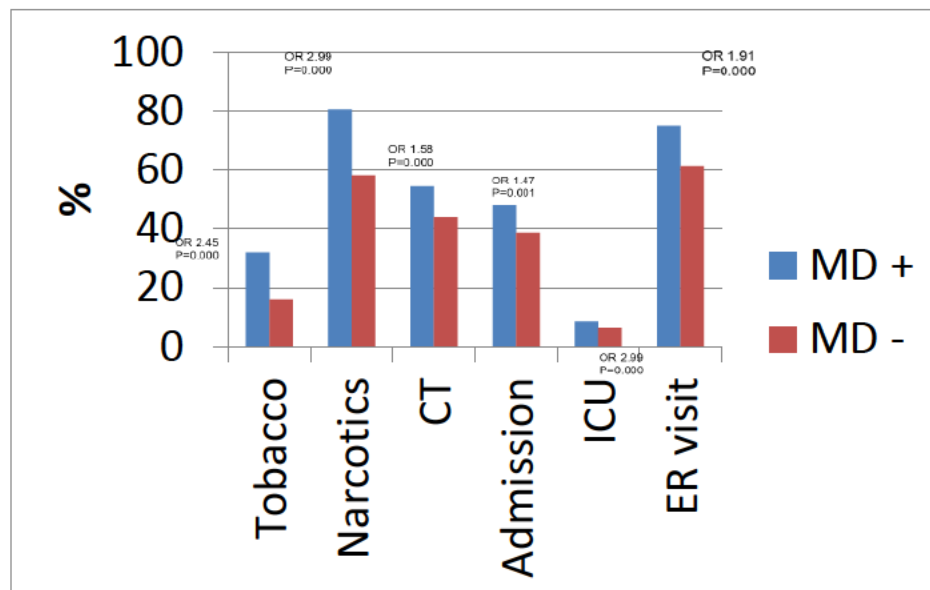


### Crohn's Disease:

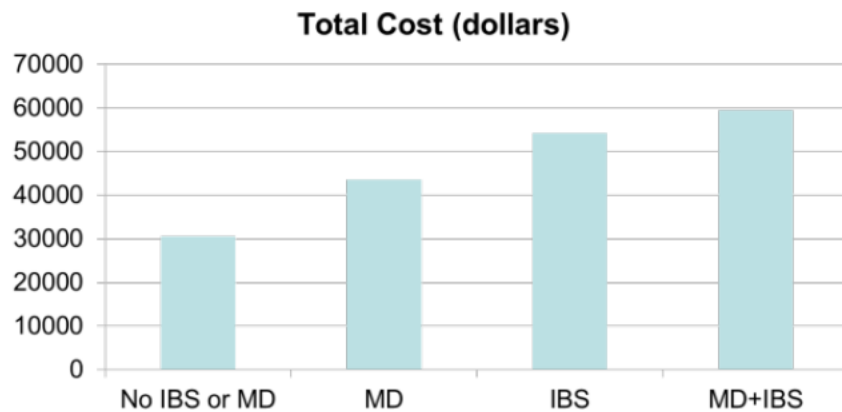
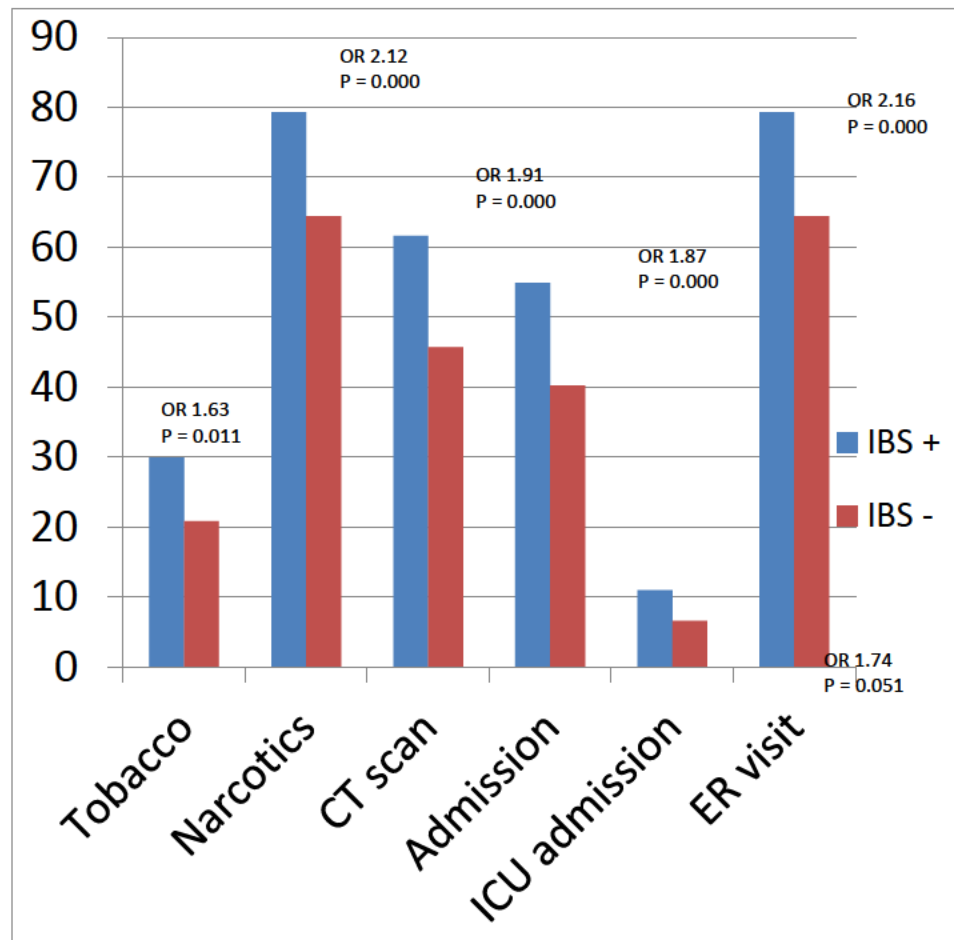
In December 2014, we presented data at the Advances in IBD conference in Orlando, FL which characterized all the new diagnoses of Crohn's disease in a nationwide military population. An incidence of 7.3/100,000 was calculated, which is similar to other recent large studies of western populations. In addition to demographic data, other important clinical data were collected on each new diagnosis. Patient comorbidities and complications of IBD (such as venous thromboembolism), as well as important clinical outcomes such as rate of surgery, inpatient admissions and ER visits were collected and compared to other large population studies. From our data we identified fairly low rates of complications of IBD and IBD treatments (CMV infection 0.3%, osteopenia/osteoporosis 6.4%, VTE 2.4%) as well as a lower rate of surgery (8.3% within the first two years of diagnosis). Admission rates (38%) and ER (64%) were high but comparable to other studies, which demonstrates the high rate of utilization of health care resources that Crohn's Disease patients pose. Another important finding in this study was the frequent use of steroids (42%) and relatively infrequent use of biologic medications (8%). Finally, a high rate of mood disorders (34%) and prescriptions for neuropsychiatric medications (50%) was discovered. It was this finding that lead to us to look more closely at those CD patients with concomitant mood disorders and irritable bowel syndrome (IBS). We found that patients with newly diagnosed CD and concomitant mood disorders or IBS were more likely to be admitted to the hospital, more likely to go to the ER, to undergo CT scans, and to be prescribed narcotics. Total medical costs were also increased for patients that had IBS, mood disorders or both, compared to patients with CD that did not have these diagnoses. This data has been accepted for a poster presentation at the upcoming Digestive Disease Week conference in Washington, DC in May 2015. See graphs below:



Crohn's patients with concomitant Mood Disorders (MD) have higher rates of tobacco use, narcotic prescriptions, more likely to undergo CT, more likely to be admitted, and more likely to visit the ER compared with Crohn's patients without MDs. They also cost more in total healthcare dollars than CD patients without MD.



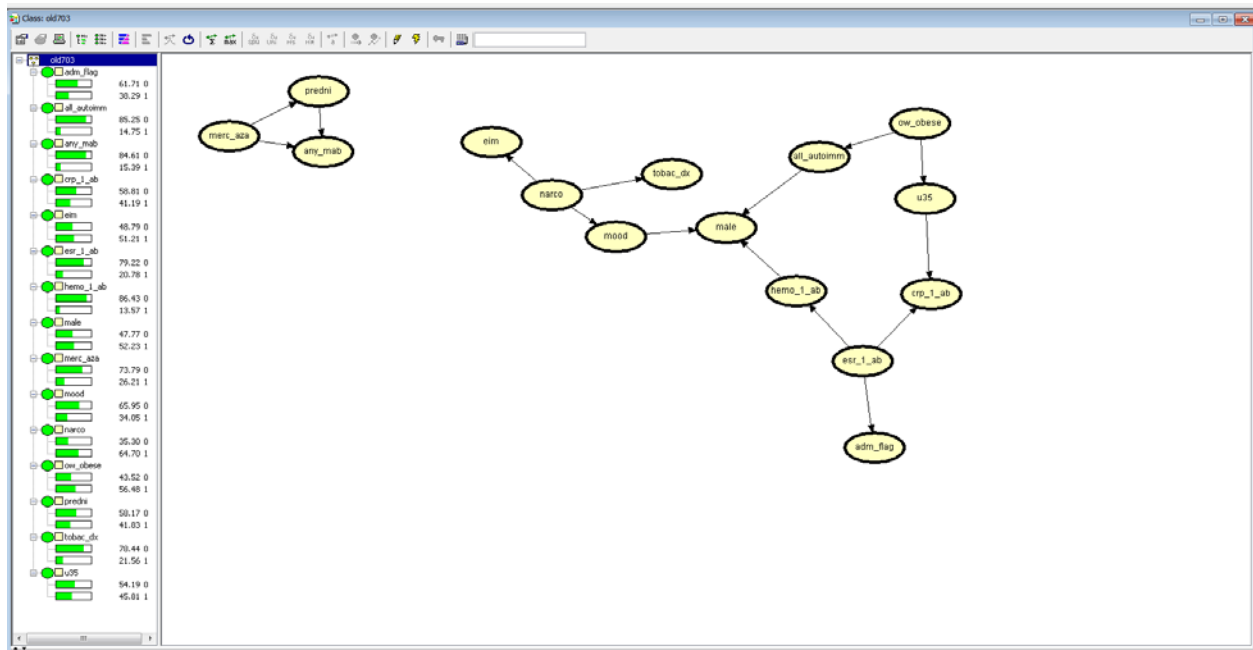
Crohn's patient's with irritable bowel syndrome also have higher rates of tobacco use, narcotic prescriptions, more likely to undergo CT, more likely to be admitted, and more likely to visit the ER compared with Crohn's patients without IBS and cost more in total health care costs.



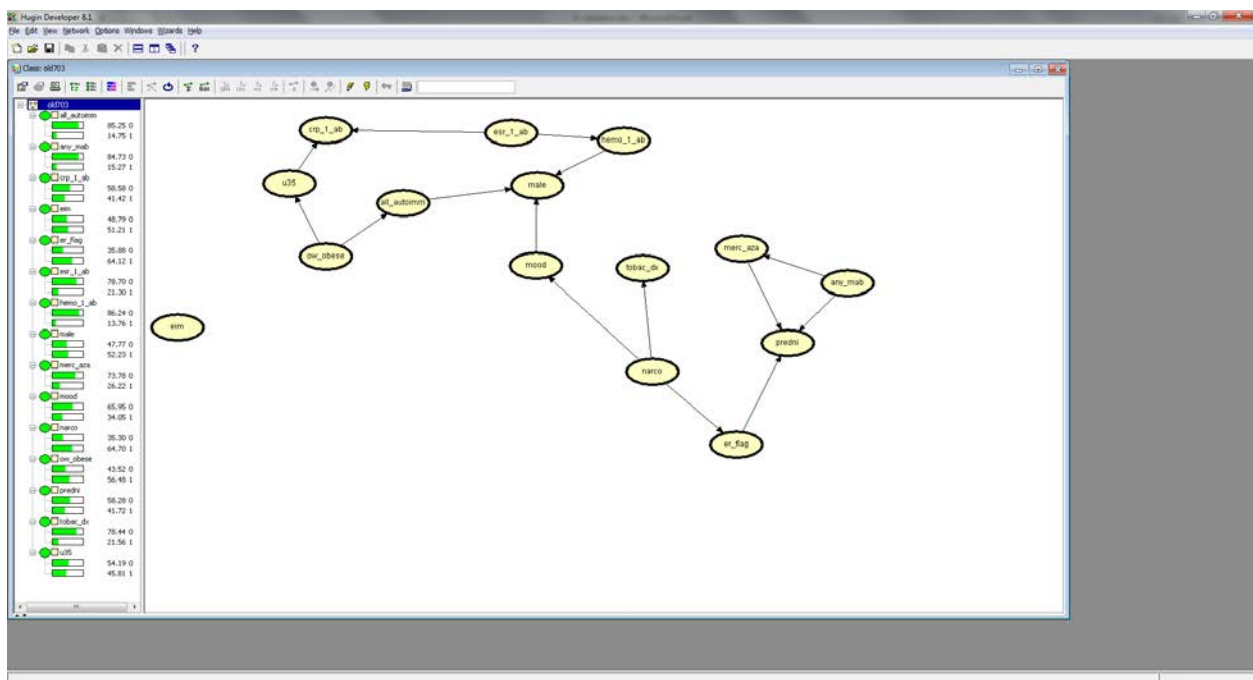
In preparation for upcoming scientific conferences (ACG 2015) we prepared the Bayesian Network analysis data showing our predictive models for CD for GI surgery, hospital admission and ER admission.



example, if a large number of patients in the catchment area are utilizing Biologic medications, placing a colo-rectal surgeon at the closest MTF would make sense from a fiscal and patient care perspective.



Regarding hospital admission, elevated serum inflammatory markers specifically ESR > 35 was shown to be independent predictor of hospital admission. However, the ROC for this model was small, (AUC of 0.53). Thus this model likely suffers from small numbers of the measured outcome and given that ESR is thought to be less useful than C-reactive protein (CRP) as a marker for disease severity is likely a poor model.



Lastly, we were able to use Bayesian Network analysis to predict ER visits in this CD cohort with a moderate level of success (ROC, AUC of 0.68). It is very interesting that narcotic use was the only independent predictor in the network analysis. It is known that narcotic use in CD patients is a risk

factor for morbidity, however in this analysis it is not known whether narcotic use is making CD worse, leading to ER visits or whether narcotic use is worsening visceral pain symptoms and though CD is relatively quiescent independently leading to ER visits.

#### **Listed Novel Research Methods**

1. Application of algorithmic computer assignment of etiology for cases of AP to a worldwide electronic medical record. This had previously only been accomplished at a single center or smaller regional network and never in the MHS.
2. Development of computer code and multisource software platform in collaboration with Kennell and UPMC to allow for extraction of important clinical variables from a digital data repository owned by the DoD utilizing existing data and through modifications to new technologies (GIANT, MEGASCOPE). To my knowledge we are the first group to begin developing working tools for this type of large data cohort development.
3. Creation of ontological dictionary specific to GI surgery and application of text concepts data extraction software to a large cohort of surgical pathology reports from a digital data repository owned by DoD with successful identification of CD related surgeries
4. Computer learning algorithmic analysis with Bayesian network analysis to create a successful predictive model of AP.
5. Identifying possible clinical predictors of severe phenotypes in CD and at the same time creating a predictive model looking for clinical predictors of the need for CD surgery.
6. Computer learning algorithmic analysis with Bayesian network analysis to create a successful predictive model of CD.

#### **4. CONCLUSION:**

Successful data extraction with preset specific case definitions and outcomes based results is establishing a model for utilizing our worldwide military clinical EMR as an effective tool for future research, resource allocation and planning with regard to complex chronic illnesses. Application of BN analysis to this robust data establishes an example of how clinical researchers can use the clinical data collected in the military EMR to provide answers to important clinical questions and empower military physicians to make better management decisions in caring for our military beneficiaries.

#### **5. REPORTABLE OUTCOMES:**

See attached Project Timeline

# Project Timetable

TASK	Y1Q1	Y1Q2	Y1Q3	Y1Q4	Y2Q1	Y2Q2	Y2Q3	Y2Q4	Y3Q1	Y3Q2	Y3Q3	Y3Q4	Status
HMJF Contract with Kennell, (Task 1)	X												Complete
Protocol Approval (Task 1)				X									Complete
Secondary Approval at MRMC-HRPO (Task 1)				X									Complete
TMA Approval at Kennell (Task 1)				X									Complete
Meeting at Univ Pittsburg (Task 1)				X									Complete
Kennell; data extraction, AP cohort					X	X							Complete
Kennell creates AP data set							X	X					Complete
<b>PROJECT EXTENTION</b>								X				X	Complete
<b>Bayesian Model for AP at WR</b>										X			<b>Complete</b>
Establish definitions for CD extraction											X		Complete
Contract with UPMC. Text Mining												X	Complete
Text Concepts data extraction												X	Complete
Kennell creates CD data set												X	On Schedule
CD BN at WR												X	On Schedule
Compare BN Models at WR and U Pitt												X	On Schedule
Present AP Data, National Meeting								X		X		X	Complete
Present CD Data, National Meeting												X	On Schedule
Full Manuscript												X	On Schedule

Green box = completed; yellow box = on schedule; CD= Crohn's disease; AP=acute pancreatitis; BN= Bayesian Network  
WR= Walter Reed; U Pitt= University of Pittsburgh

**6. PUBLICATIONS, ABSTRACTS, AND PRESENTATIONS: ( Abstracts, see appendix)**

**7. APPENDICES:**

ACG 2013, ACG 2014, DDW 2013, DDW 2015, CCFA 2014



The Epidemiology of Acute Pancreatitis within a Nationwide Military Population

Scott Cunningham, MD, Corrine Maydonovich, BS, Ganesh Veerappan, MD, John Betteridge, MD,

Walter Reed National Military Medical Center, Bethesda MD

Purpose:

The aim of this study is to examine the epidemiology of acute pancreatitis (AP) in our military health care system, and explore the demographic risk factors for the various etiologic subtypes of AP.

Methods:

Using the electronic medical records (EMR) and healthcare claims databases under the Military Healthcare System Data Repository (MDR), the total number of admissions, laboratory data and clinical outcomes related to acute pancreatitis (ICD9-CM 577.0) were examined from October 1, 2008 to September 30, 2012. An established and published hierarchical methodology using ICD-9 codes was used to assign etiology for each pancreatitis case. Once etiologies were established, we used other raw data and outcomes to better characterize each etiology category.

Results:

Over this 4-year period 3134 cases of AP were identified in a study population of 2,973,523 patients. The cumulative incidence of AP was 26 per 100,000 patients per year. Patients with AP had a mean age of  $47 \pm 19$  years, 53% male, and 53% Caucasian. Incidence rates were similar among genders (male to female, RR 0.96) and races (Caucasian to African Americans, RR 1.12). The overall mortality rate was 0.6% (20/3134). Six percent of patients had severe AP as defined by an ICU stay of >48 hours. Idiopathic pancreatitis was the most common single cause of AP (45.3%), followed by gallstone pancreatitis (27%). Alcohol accounted for 16% of the cases of AP. Blacks were nearly twice as likely as Whites to have alcohol as the cause of pancreatitis (OR 1.99). Whites were nearly twice as likely as Blacks to have gallstones as the cause of their pancreatitis (OR 2.01). Females were significantly more likely to have biliary pancreatitis than males (OR 1.95). Males were significantly more likely to have alcoholic pancreatitis (OR 3.39). The severity of pancreatitis did not differ significantly among different etiologies. Alcohol was the most common cause of recurrent pancreatitis, ranging from 40% of the second admissions; to more than 80% of 6<sup>th</sup>, 7<sup>th</sup>, and 8<sup>th</sup> admissions.

Conclusions:

Our study provides an assessment of the burden of AP within a diverse, nationwide population. A hierarchical method in a vast EMR reliably allowed an accurate assignment of etiology. Our mortality rate was significantly lower than previous studies, possibly due to our relatively younger and healthier population. Using this method of determining etiology, we identified idiopathic as the most likely cause of pancreatitis, followed by gallstone and alcohol. Alcohol occurred more often in men and Blacks, while gallstone pancreatitis was more likely in Whites and women. Interestingly, alcohol disease was the most common cause of recurrent pancreatitis.

## The Epidemiology of Acute Pancreatitis within a Nationwide Military Population

Scott Cunningham, Corrine Maydonovich, Ganesh Veerappan

Walter Reed National Military Medical Center

### Introduction:

Acute pancreatitis (AP) is a commonly encountered gastrointestinal illness, with a broad clinical spectrum, ranging in severity from mild to life-threatening. Due to this clinical heterogeneity, it is important to identify high risk patients at presentation to improve patient care while optimizing clinical resources. Numerous population-based studies describe demographic data, in an attempt to identify risk factors that predispose to acute pancreatitis. Scoring systems have developed based on various clinical, radiological and laboratory data with reasonable success.

### Aim:

The aim of this study is to examine the epidemiology of acute pancreatitis in our military health care system, and explore demographic and laboratory factors associated with a more severe course of pancreatitis.

### Methods:

Using the electronic medical records and healthcare claims databases under the Military Healthcare System Data Repository (MDR), the total number of admissions and outcomes related to acute pancreatitis (ICD9-CM 577.0) were examined over a period from October 1, 2008 to September 30, 2012. Data extracted for this analysis included demographic (age, gender, race, body mass index) and specified laboratory data that may be predictive of severe pancreatitis. For analysis, lab values were dichotomized based on prior scoring systems to predict severity (Creatinine > 2.0 mg/dL, white blood cell count (WBC) > 16,000 cells/ $\mu$ L, glucose > 200 mg/dL, lactate dehydrogenase (LDH) > 350 IU/L, aspartate aminotransferase (AST) > 250 IU/L, calcium < 8 mg/dL, hematocrit 44%, blood urea nitrogen (BUN) > 25 mg/dL). Outcomes included mortality, ICU admission > 48 hours, and evidence of end organ damage. Severe pancreatitis was defined as an ICU admission > 48 hours. Logistic regression was performed to identify independent factors associated with severe pancreatitis.

### Results:

Over this 4 year period, 2,197 cases of acute pancreatitis were identified in a study population of 536,929 patients. Patients with acute pancreatitis had a mean age of  $41y \pm 14$ , 49% male, and 53% Caucasian. The cumulative incidence of AP was 1.0 per 1000 patients per year. Acute pancreatitis consisted of 0.4% of all hospitalizations over this period. Incidence rates were similar among genders (male 1.04/1000, female 1.01/1000) and races (Caucasians 2.33/1000, African Americans 2.13/1000). The overall mortality rate was 0.2%. Five percent (107/2,197) of patients had severe pancreatitis.

## 2013 DDW Abstract

Independent predictors of severe pancreatitis include age > 55 (OR, 1.6; P< 0.04), male gender (OR, 1.5; P<0.05), white blood cell count >16,000 cells/mm<sup>3</sup> (OR, 2.7; P <0.01), serum glucose >200 mg/dL (OR, 2.1; P<0.02) and serum calcium <8 mg/dL (OR, 6.3; P<0.01).

### Conclusions:

Compared to other recent population-based epidemiologic studies, we report a higher incidence of AP (1.0 vs. 0.6 per 1,000), less severe pancreatitis (5% severity) and lower mortality (0.2%). Independent predictors of severe pancreatitis include age >55, male gender, elevated wbc, elevated glucose and decreased serum calcium, which have been validated in other studies.

The Use of the Bayesian Analysis to Predict Severe Acute Pancreatitis in a Nationwide Military Population

Scott Cunningham, MD, Corinne Maydonovich, BS, John Betteridge, MD,

Walter Reed National Military Medical Center, Bethesda MD

Purpose:

Predicting clinical outcomes is a valuable asset in improving patient care and optimizing medical resources. Acute pancreatitis (AP) is a relatively common gastrointestinal disease which can vary widely in its clinical severity. Bayesian network analysis is a method which may be useful to predict clinical outcomes of a variety of medical conditions. In this study, we used Bayesian network analysis to identify predictors of severe acute pancreatitis within our military health care system.

Methods:

Using data derived from the Military Healthcare System Data Repository (MDR), the total number of admissions, laboratory data and clinical outcomes related to AP were examined from October 1, 2008 to September 30, 2012. Cases of AP were identified using ICD-9 codes associated with hospital admissions. Using a definition of a severe course as requiring ICU stay > 48 hours, a Bayesian network analysis was applied to this cohort. Models were created using 80% of the data set then tested for accuracy with the remaining 20% of data. Multivariate logistic regression analysis was then performed to confirm these results.

Results:

A total of 3,134 cases of acute pancreatitis were identified out of a total study population of 2,973,523 for a cumulative incidence of 26 per 100,000. Six percent of these admissions (194 cases) were classified as severe AP. Bayesian network analysis identified 4 predictors of severe AP. The predictors included white blood cell count (WBC) >16,000 cells/mL, aspartate aminotransferase (AST) >250 IU/mL, serum creatinine (SCr) >2mg/dL, serum calcium (Ca) <8 mg/dL OR 4.15. These same 4 predictors were then confirmed by multivariate logistic regression analysis. When applied to the data set, the accuracy of the Bayesian model was 70.4%.

Conclusions:

Bayesian network analysis identified 4 clinical variables as predictors of severe AP and these were confirmed by a logistic regression analysis. Despite their agreement, the Bayesian network model had modest accuracy. Our results were limited the inability to obtain certain important clinical data pancreatitis (i.e radiographic findings, vital signs, mental status) from the electronic medical record, which have been established as clinically relevant predictors of AP. Despite these limitations, our results do indicate that Bayesian analysis is useful in predicting variables for a defined outcome. With a more complete data set, the accuracy of the Bayesian model will likely improve.

The Incidence of Crohn's Disease and Rate of Early Surgery in the U.S. Military Health Care System

Scott Cunningham, MD, Corrine Maydonovich, BS, John Betteridge, MD,

Walter Reed National Military Medical Center, Bethesda MD

Background:

Crohn's Disease (CD) is chronic, incurable disorder that can have a variable clinical course. CD patients can have multiple complications as well as significant comorbidities. Surgery is sometimes required to treat complications of CD such as fistulas or strictures, as well as to remove affected portions of the GI tract. The need for surgical intervention may be decreasing with the advent of more effective medical therapy for CD. The aim of this study was to calculate the incidence of CD in a nationwide military population and to further characterize the CD population, while also assessing outcomes such as surgery, hospitalization and emergency room visits.

Methods:

Using the electronic medical records (EMR) and healthcare claims databases under the Military Healthcare System Data Repository (MDR), the total number of new diagnoses of Crohn's disease were examined 2009-2012. The data collected were derived from a nationwide network of military treatment facilities and included a total study population of 3.2 million. For each new diagnosis of CD, demographic variables, basic laboratory data, other chronic comorbidities, medications used to treat CD, and complications (such as venous thromboembolism) were analyzed. Outcomes such as ER visits, admissions, surgeries and deaths were also recorded up to two years from the time of diagnosis.

Results:

Over the three-year period, 703 new diagnoses of Crohn's disease were identified. The average annual incidence of CD was 7.3 per 100,000 patients per year. Fifty-two percent of patients were male. Fifty-three percent of patients were under the age of 35. The average body mass index at diagnosis was 26. Twenty two percent reported tobacco use. Within two years of diagnosis, 9% patients underwent surgery for CD. Within two years of diagnosis, 38% of patients were admitted to the hospital (71% of these admissions were directly related to CD). Seven percent of patients were admitted to the ICU. Sixty-four percent were seen in the ER for any reason during the study period. Two percent required TPN. Seven percent of patients were treated with biologic therapy. Forty-two percent of patients were treated with corticosteroids. Thirty one percent of patients were treated with immunomodulator therapy. No patients died. Thirty-four percent of patients were diagnosed with a concomitant mood disorder and 50% of patients were prescribed a neuropsychiatric medication. Twelve percent of patients were also diagnosed with irritable bowel syndrome. Three percent of patients were diagnosed with *Clostridium difficile* colitis. Two percent of patients had a pulmonary embolism or deep vein thrombosis. The mean cost treat each patient was \$33,057.24.

Conclusions:

Our study shows a similar incidence rate of Crohn's disease compared to another earlier large population studies in the US (7.3/100,000 compared to 7.9/100,000), which suggests a stable incidence over time. The early surgery rate found in our study is lower than that found in other published studies, although this correlates with a trend of decreasing need for surgery in the era of biologic medications. Our hospitalization rate is higher than that reported in an earlier study in Olmsted County, but lower than that reported in a population study from Denmark. Our data show a relatively low percentage of patients who were treated with biologic therapy. Also notable is the high rate of concomitant mood disorders and use of neuropsychiatric medications. (will try to find some data to compare our numbers to re: mood disorders/psych meds).

. (Denmark, 2003-2005: 12% resection within first year; Norway 14% risk of surgery within first year; Olmsted County, 1940-2001: 24% risk in first year.)

. (Copenhagen County, 1962-1987: 83% admitted within first year; Canada, 1994-2001: 25% of all Crohn's patients were admitted; Olmsted County, 1970-1999: 69% of patients hospitalized at least once; cumulative risk of hospitalization at one year 32%)

Background:

Crohn's Disease (CD) is chronic, incurable disorder that can have a variable clinical course. These patients can develop multiple complications from their disease. In addition, the treatments for CD (both medical and surgical) can impose serious morbidity. The medical costs for treatment of CD can be great, owing to the expensive medical therapy, frequent use of medical resources including inpatient admission. The purpose of this study was to calculate the incidence of CD in a nationwide military population and to explore the demographic variables, evaluate for the presence of comorbidities, complications and to evaluate the need for surgical intervention within the first two years of diagnosis.

The treatment for CD can sometimes involve very expensive medications with potentially serious side effects, as well as surgical procedures with high morbidity. It is beneficial to both patient and to the health care system to be able to identify the patients who are more likely to experience a more severe or complicated disease course, so that they may receive more aggressive care early, and to spare such interventions on patients who may not need them. The aim of this study is to examine the incidence of (CD) in our military health care system and to characterize the severity of the

The Prevalence and Impact of Mood Disorders and Irritable Bowel Syndrome in a Nationwide Military Crohn's Disease Population

Scott Cunningham, Corrine Maydonovich, Ganesh Veerappan, John Betteridge

Walter Reed National Military Medical Center

Introduction:

Patients with Crohn's Disease (CD) have been observed to have higher rates of mood disorders (MD) compared to the general population. Irritable bowel syndrome (IBS) is a common, but benign, gastrointestinal disorder whose symptoms can mimic that of CD. Patients with IBS have also been observed to have higher rates of mood disorders than the general population. Patients with CD and IBS can pose clinical challenges as it can be difficult to differentiate symptoms of active CD from IBS symptoms.

Aim:

To examine the prevalence of IBS and MD in a newly diagnosed population of CD patients, and to examine how these diagnoses affect the treatment and utilization of healthcare resources compared to patients with CD without IBS or MD.

Methods:

Using the electronic medical records and healthcare claims databases under the Military Healthcare System Data Repository (MDR), the total number of new diagnoses of CD were examined over a period from 2008-2012 using ICD-9 codes. The prevalence of MD and IBS was calculated within this population, also using ICD-9 codes. Demographic data, laboratory data, prescription history and outcomes such as ER visits and hospitalizations were collected on all the CD patients during this time period.

Results:

The prevalence of MD and IBS within the CD cohort was 35.7% and 12.5% respectively. Patients with IBS were more likely to visit the ER (OR 2.12, CI 1.42-3.15), to be admitted (OR 1.81, CI 1.3-2.51), to undergo computed tomography scans (CTs) (OR 1.901, CI 1.36-2.66) and to be prescribed narcotics (OR 2.12, OR 1.42-3.15). Patients with MDD were also more likely to visit the ER (OR 1.91, CI 1.48-2.45), to be admitted (OR 1.47, CI 1.17-1.84), to receive CT scans (OR 1.53, CI 1.22-1.92) and to be prescribed narcotics (OR 2.99, CI 2.29-3.90). Rates of surgery and use of biologic medications and TPN were not more frequent in these patient populations. Patients with MD and IBS had higher total costs of care over the study period compared to those who did not have these diagnoses.

Conclusions:

Our study shows a high rate of MD within a newly diagnosed CD cohort in a nationwide military population. The prevalence of IBS in the CD population was not increased, but likely reflects the difficulty in making the diagnosis of concomitant IBS in a newly diagnosed CD patient. Patients with IBS and MD had higher rates of hospitalization, ER visits and total costs of care, but similar rates of surgery and need for biologic therapy, which suggests that they utilize more resources but do not necessarily have more severe disease. Recognition and treatment of MD and IBS in CD patients may help reduce unnecessary admissions, ER visits and CT scans.